

Metodi Montecarlo per la Fisica Sperimentale

Sergio Giudici

Dipartimento di Fisica – Università di Pisa

sergio.giudici@unipi.it

Informazioni sul corso

- **Codice 185BB ; Settore FIS-01 ; 6 CFU , 36h**
- **Obiettivi:** Students are expected to acquire The mathematical basis of statistical sampling as a tool for computational problem. Skillness in applying pseudo-random number techniques for simulation in physics. Skill in the realization of algorithms based on pseudo-random number
- **Modalità Esame:** Final oral exam and Written report. During the exam student must be able to demonstrate to know the mathematical basis of the statistical sampling method (50 % weight) and must be able to present and discuss with property of expression a specific problem (50% weight) assigned by the teacher, considering the interests of the student. Student must demonstrate the ability to approach a circumscribed research problem and organize an effective exposition of the results.

Programma del corso

- **Richiami di statistica e Fondamenti teorici del Metodo**
- **Numeri pseudo-casuali e campionamento di distribuzioni assegnate**
- **Architettura di una simulazione: generazione e ricostruzione**

Sono discusse applicazioni specifiche tenendo conto degli interessi dei partecipanti. Esempi di argomenti trattati:

- **Funzione di risposta di un detector in Fisica delle Alte Energie**
- **fit cinematici e algoritmi di ricostruzione**
- **simulazione della interazione radiazione-materia**
- **simulazioni Hamiltoniane con effetti stocastici**

Materiale didattico:

- **Testo di riferimento: Rotondi A. et al, «Probabilità, Statistica e Simulazioni» , Springer 2004 (I capitoli relativi alla simulazione)**
- Altro materiale utilizzato in passato
- Levin C.S. "Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution", Phys. Med. Biol. 44 (1999) 781–799
- Pia M.G. et al. "Epistemic and systematic uncertainties in Montecarlo Simulation: an investigation in proton Bragg peak simulation", arXiv , 2010
- Beisbart C., How Can Computer Simulations Produce New Knowledge?, European Journal for Philosophy of Science 2 (2012), S. 395-434

Esempi di Studi presentati al colloquio finale

- Studio della risoluzione spaziale del sistema GPS ed effetti dovuti alla ionosfera
- Simulazione di uno spettrometro per un esperimento di Fisica delle Alte Energie
- Simulazione della tracciatura PET di un fascio di protoni incidente su target d'acqua
- Tecniche di ricostruzione di campo stellare tramite l'algoritmo STARFIND

Storia (breve) dei Metodi MC

- **1733** : Problema dell'ago di Buffon «**Essai d'arithmétique morale**»
- **1940s** : L'invenzione del metodo è attribuita a S. Ulam, J. Von Neumann, N. Metropolis. I primi due erano coinvolti nel Progetto Manhattan (bomba atomica). Si pensi al problema di determinare la massa critica di Uranio affinché si inneschi la reazione a catena. (cfr. Markov Chain)
- Ulam, von Neumann e Metropolis hanno sviluppato algoritmi per implementare il campionamento statistico su calcolatori e studiato modi per facilitare la soluzione di problemi deterministici (Integrazione numerica, eq. alle derivate parziali, etc..) utilizzando metodi basati sul campionamento statistico.
- **1949**: Ulam e Metropolis, «*The Montecarlo Method*», Journal of American Statistical Association, 44 (247) , 335-341, 1949
- Le applicazioni del Metodo sono ora vastissime e hanno seguito di pari passo l'aumento delle prestazioni dei calcolatori.

Ago di Buffon (1733)

Su un piano decorativo a strisce parallele (per esempio un pavimento in parquet o un tappeto a strisce), tutte della stessa larghezza, si lancia a caso un ago. Qual è la probabilità che l'ago cada in una posizione in cui interseca il confine tra due strisce?

Se la spaziatura è pari alla lunghezza dell'ago si ha $p = \frac{2}{\pi}$ (dimostrarlo per esercizio)

Indicato con N_s il numero di successi su un totale di N lanci si ha

$$\lim_{N \rightarrow \infty} \frac{N_s}{N} = \frac{2}{\pi}$$

Da cui segue la stima «pi greco» (legge grandi numeri)

$$\frac{N_s}{N} \approx \frac{2}{\pi}$$

Il problema è valutare la bontà della approssimazione ovvero stimare in senso probabilistico

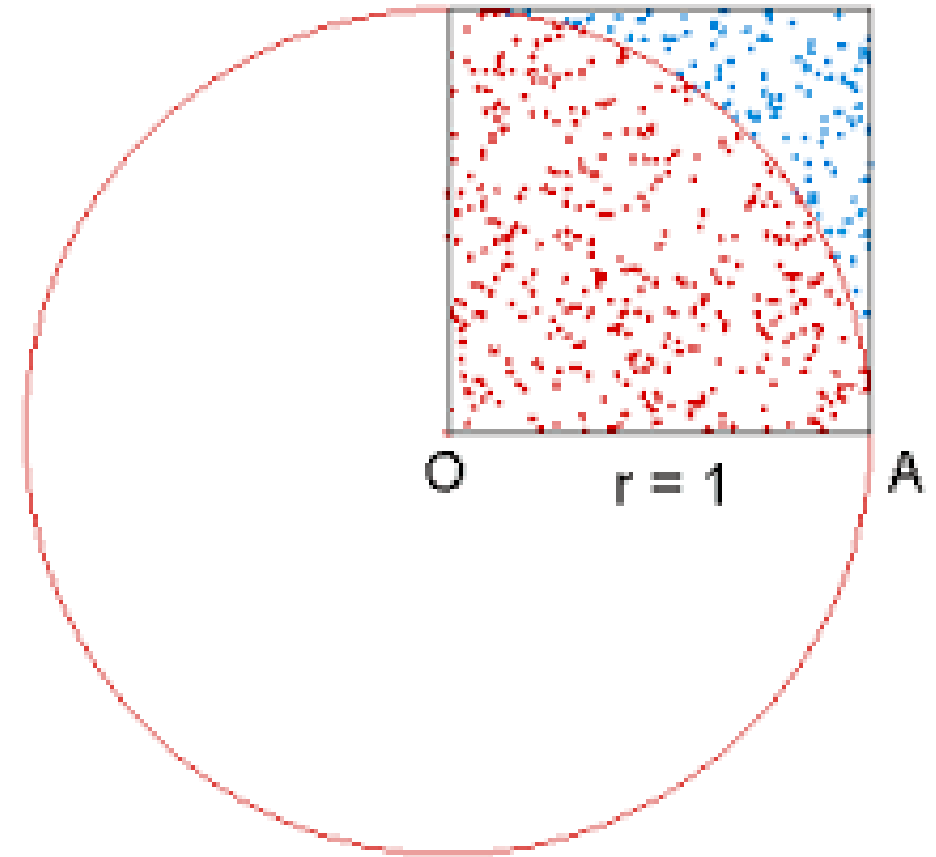
$$\left| \frac{N_s}{N} - \frac{2}{\pi} \right| < ?$$

Una altra stima di «pi greco»

Se x e y sono aleatorie uniformi in $[0,1]$

$$P(x^2 + y^2 < 1) = \frac{\pi}{4}$$

$$\lim_{N \rightarrow \infty} \frac{N_s}{N} = \frac{\pi}{4}$$



Stima dell'errore

- Il numero di successi N_s ha evidentemente un carattere «binomiale»
- Con varianza $\Delta N_s = \sqrt{\frac{\pi}{4} \left(1 - \frac{\pi}{4}\right) N}$
- Se N è abbastanza grande in modo che la binomiale «assomigli» alla Gaussiana, abbiamo con un livello di confidenza ~68%

$$\left| \frac{N_s}{N} - \frac{\pi}{4} \right| < \frac{\Delta N_s}{N} \approx \frac{0.41}{\sqrt{N}}$$

Esercizi

- Discutere l'errore nel caso in cui si stimi π greco con un ottante di sfera ovvero con x, y, z uniformi in $[0, 1]$ e valutando la probabilità

$$P(x^2 + y^2 + z^2 < 1)$$

- Si consideri il caso con n variabili uniformi in $[0, 1]$ (n -sfera) e dire se «conviene» spingersi a dimensioni elevate

Numeri casuali uniformi in $[0,1]$

- La stima di π si basa sul campionamento uniforme dell'intervallo $[0,1]$
- Come si realizza tale uniformità ?
- Si consideri l'intero I a caso in formato binario con 64 bit

$I = 01101010010101001...00111$

- **Si tratta di una successione casuale di 0 e 1 che in linea di principio si potrebbe generare tirando una moneta 64 volte.**
- Il massimo di questi interi è **$IMAX = 2^{n-1}$**
- **Allora $I/IMAX$ è uniforme in $[0,1]$**

Implementazione Hardware di rand()

- Cosa potrebbe essere l'analogo di «tirare un moneta» all'interno di un calcolatore? Come si fa ad essere sicuri della uniformità ?
- Alcuni hanno proposto come generatore hardware di rand() le fluttuazioni della corrente che circola in una resistenza: Un amperometro (abbastanza sensibile) sarebbe un generatore di numeri casuali
- In fisica delle alte energie si usa qualche volta un processo casuale per sorteggiare l'istante in cui «scaricare» lo stato di un detector (un trucco per studiare l'attività accidentale come raggi cosmici, interazione del fascio con gas residuo, canali noisy) . Per «triggerare» l'istante in cui fare il download si può usare il decadimento radioattivo come avviene nell'esperimento mentale del «gatto di Schroedinger».
- L'emissione di una particella alfa da parte di nucleo radioattivo è, di fatto, un generatore di numeri casuali!

Sequenze non ripetibili

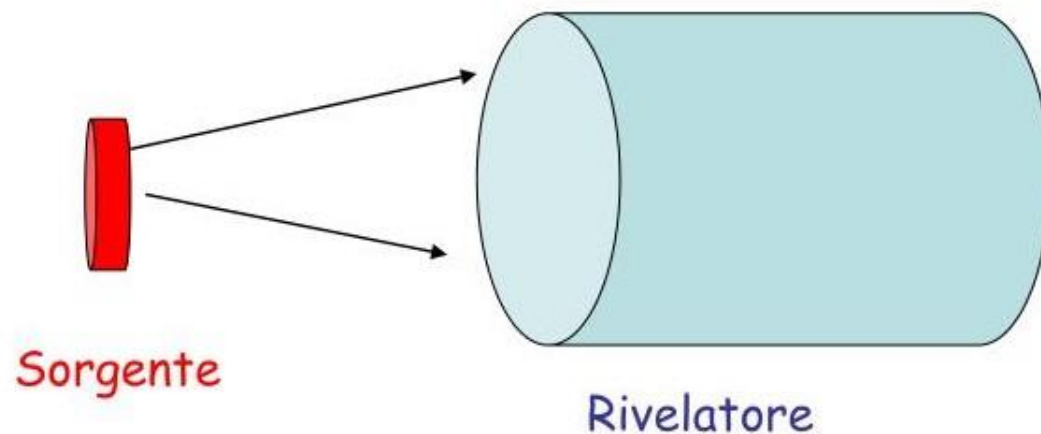
- La soluzione amperometro o emissione radioattiva garantisce l'uniformità e fornisce una sequenza di numeri casuali tuttavia
- A meno che non si «salvi» la sequenza, non sarà ripetibile.
- Perché vogliamo garantire la ripetibilità ?
- L'esigenza nasce dal caso delle «Simulazioni Correlate»

Esempio: Accettanza geometrica

Angolo solido sotteso da un cilindro piano posto ad una distanza r da una sorgente estesa

In generale il calcolo dell'angolo solido comporta una integrazione delle funzioni di Bessel mediante metodi numerici.

Per una soluzione approssimata vedi Knoll, Cap.4



Simulazioni Correlate

- Invece di tentare l'integrazione numerica si possono «generare» eventi (nel senso Montecarlo!) e stimare l'accettanza per via statistica
- In fase di progettazione si confrontano due diverse geometrie per il detector e si ottengono accettanze diverse: $A1$ e $A2$
- **Domanda?** $A1$ e $A2$ differiscono perché la geometria è cambiata oppure a causa delle fluttuazioni statistiche nella generazione degli eventi? Come si possono «separare» i due effetti?
- Risposta: Basta «illuminare» le due diverse geometrie con gli stessi eventi: la differenza tra $A1$ e $A2$ sarà dovuta solo al cambio di geometria!
- **Per garantire la medesima «illuminazione» bisogna ripetere la stessa sequenza di numeri casuali → ripetibilità**

Liste di numeri Pseudo-Casuali - Metodo delle Congruenze lineari

vedi https://it.wikipedia.org/wiki/Generatore_lineare_congruenziale

Si consideri la successione definita per ricorrenza

$$X_{n+1} = (aX_n + c) \bmod m$$

I numeri cadono nell'intervallo $0 \leq x \leq m-1$ se dividiamo per $m-1$ fornisce numeri in $[0,1]$.

La successione è deterministica ma «sembra» casuale. Provate a fare l'esercizio di ingegneria inversa: data la sequenza determinare a, c, m ... (ma anche solo immaginare che dietro ci sta la congruenza lineare!)

Gioco: Dire come prosegue la successione 3,3,3,7,6,3,5,4,4,5, (soluzione a fine corso)

In un episodio della serie Doctor Who si scopre di vivere in mondo virtuale (à la Matrix) ... generando a caso numeri interi. Trovare l'episodio ... non ricordo il titolo!

Sequenze pseudo-casuali e Periodicità

- La lista di pseudo-casuali è ripercorribile: basta partire dallo stesso **seme X_0**
- Esiste letteratura copiosa sui numeri pseudo-casuali (teoria dei numeri) ... ovviamente non ce ne occupiamo !
- Le congruenze lineari forniscono liste che al massimo hanno periodo uguale a m (dimostrarlo!)
- Ovviamente sono apprezzati gli algoritmi che hanno il periodo maggiore possibile (Fare una ricerca sui generatori random)
- Esercizio: **Dire cosa succede se stimate «pi greco» utilizzando n volte la stessa successione di m numeri pseudo-casuali**

Montecarlo VS Metodi analitici

Una delle applicazioni dei numeri pseudo-casuali consiste nella determinazione per via numerica di Integrali . La stima di π tirando «punti a caso» è un modo (forse non il più furbo!) per risolvere numericamente l'integrale

$$\int_0^1 \sqrt{1 - x^2} = \frac{\pi}{4}$$

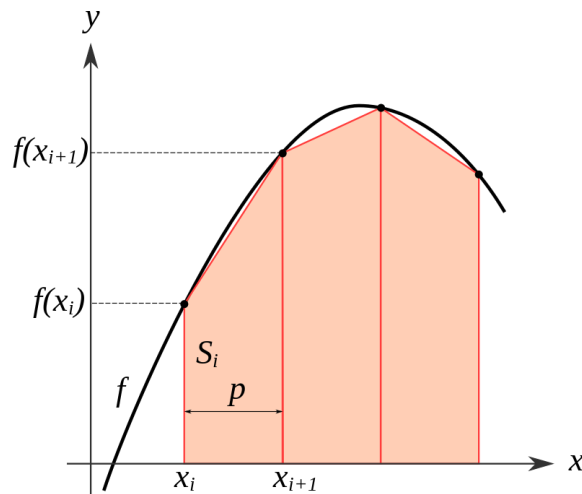
Se N indica il numero di punti sorteggiati ovvero il numero di volte che si è calcolato il valore dell'integrando, il metodo fornisce l'integrale con una accuratezza che scala come $N^{-1/2}$

Domanda: MMC è competitivo quanto a accuratezza rispetto ad altri metodi ? Vale la pena di adottarlo? ... vediamo gli altri metodi...

Metodo dei trapezio

- Un metodo classico per valutare numericamente integrali definiti consiste nel dividere l'intervallo di integrazione $[a,b]$ in N sottointervalli di uguale ampiezza $(b-a)/N$, nei quali la funzione si approssima con il suo valore medio (Regola del trapezio): L'errore che si commette ad approssimare l'area sottesa con la successione di trapezi è data da

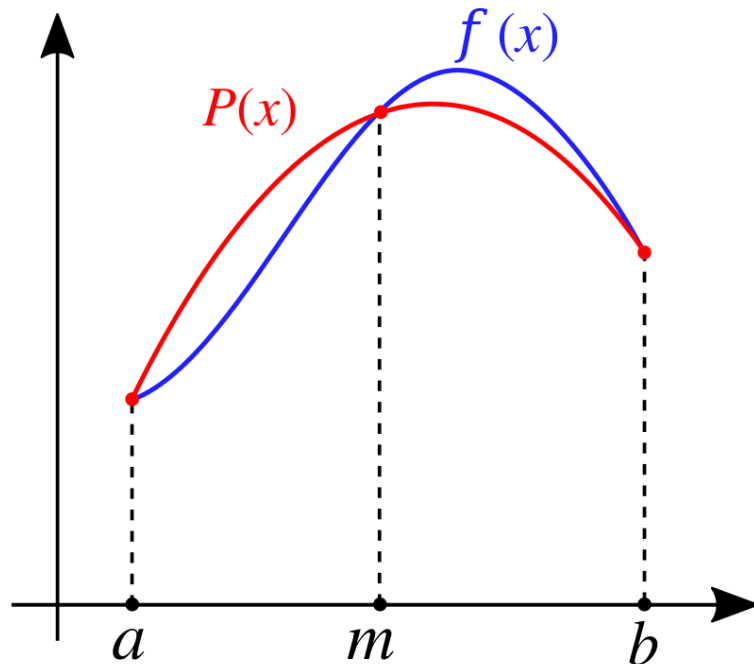
$$Error = -\frac{(b-a)^3}{N^2} f''(c) + O(N^{-3})$$



Dove c è un qualche punto in $[a,b]$. La dimostrazione segue dal teorema della media di Lagrange. [Cercare una dimostrazione]

Approssimazione parabolica

- Si possono usare altre approssimazioni. Il metodo Cavalieri-Simpson usa una approssimazione parabolica



$$Error = -\frac{(b-a)^3}{N^3} f'''(c) + O(N^{-4})$$

Dove c è un qualche punto in $[a,b]$.

Aumentando il grado del polinomio l'accuratezza migliora ma si paga il prezzo «computazionale» di dover calcolare derivate di ordine elevato !

Confronto con accuratezza MC e Integrali in dimensione d

- Il metodo dei trapezi si può estendere al caso di Integrali doppi ($d=2$)
- IL numero N di «chiamate» alla funzione integranda viene però distribuito su una griglia bidimensionale $N_x \times N_y = N$
- La precisione del metodo trapezio si abbassa di grado $N^{-2/d}$
- Invece, il metodo MC mantiene il suo carattere binomiale e qualunque sia d ha sempre la stessa accuratezza $N^{-1/2}$
- **Se $d > 4$ Montecarlo «vince» sui Trapezi**

Un teorema utile per il campionamento

- Teorema: La distribuzione della cumulativa di una PDF $dP(x)/dx$ è uniforme in $[0,1]$

$$r(x) = \int_{-\infty}^x P(u) du$$

- $P(u) > 0$ per ogni u perché $P(u)$ è una densità di probabilità $\rightarrow r(x)$ è monotona crescente.
- La probabilità di avere $r_1 < r < r_2$ è uguale a

$$\int_{x_1}^{x_2} P(u) du = r(x_2) - r(x_1) = r_2 - r_1$$

Essendo $P(r_1 < r < r_2) = (r_2 - r_1)$ allora r è uniforme in $[0,1]$

Campionamento Esatto

- Se si sa invertire l'equazione $r = \int_{-\infty}^x P(u) du$
- Allora si può sorteggiare r uniforme in $[0,1]$ e determinare x
- In questo modo si è certi che x si distribuisce come $dP(x)/dx$

Metodo del rigetto (Von Neumann, 1951)

Supponete di voler generare random una variabile x distribuita secondo una distribuzione di probabilità $P(x)$. Se il campionamento esatto non è applicabile (non si sa invertire !) allora si può utilizzare la tecnica un metodo “hit or miss” (applicabile a ogni $P(x)$ su dominio finito)

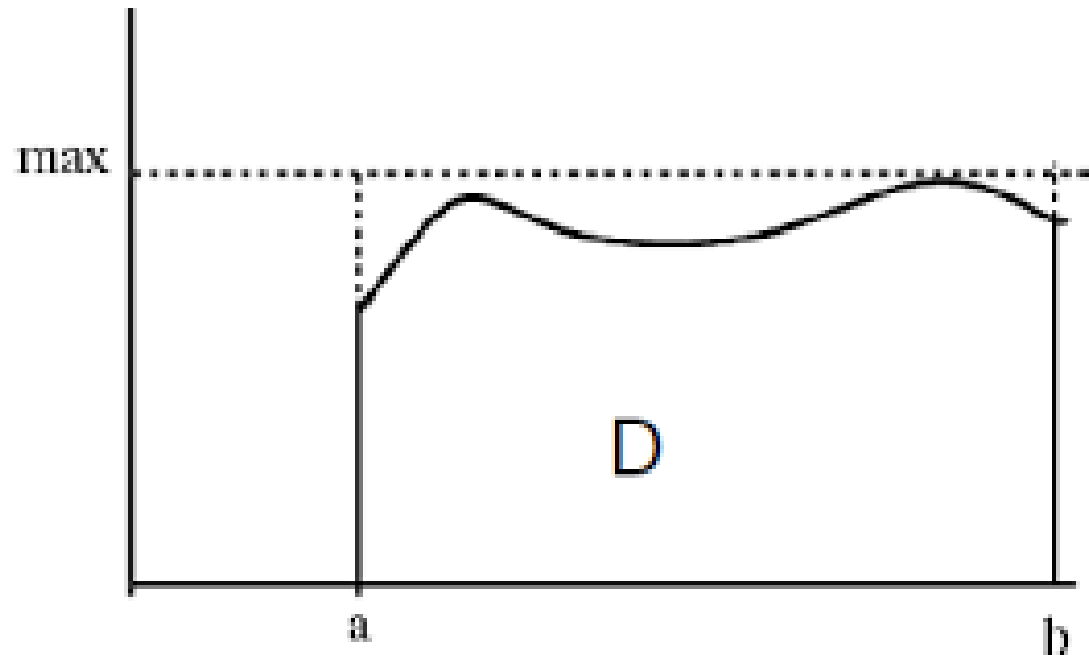
Generate due variabili casuali x, y distribuite uniformemente nei domini

$x = \text{random in } [a, b]$; $y = \text{random in } [0, \text{max}]$

x è accetta se $y < P(x)$

Altrimenti si ripete il procedimento

Fino a che non si accetta la x !



Rigetto ottimizzato

- a) determinate un'altra PDF $g(x)$ per la quale la generazione di punti casuali è semplice e veloce (metodo dell'inversione) e tale che $P(x) < g(x)$
- b) Generate x distribuita secondo $g(x)$ ed estraete un numero casuale y uniforme in $[0, g(x)]$
- c) Se $y < P(x)$ l'estrazione viene accettata, altrimenti viene rigettata e si ritorna al punto b).

Esercizi

- 1) Campionare la distribuzione $P(x) = \exp(1/(1+x^2)) - 1$
- 2) Campionare una gaussiana standardizzata sfruttando il teorema del limite centrale [Suggerimento: Se x_1, x_2, \dots, x_n sono uniformi in $[0,1]$
Come si distribuisce la media $z=(x_1+x_2+x_3+\dots+x_n)/n$?
- 3) Dimostrare che (in coordinate polari) si può determinare analiticamente la cumulativa del prodotto di due Gaussiane indipendenti $G(x)*G(y)$
- 4) Dimostrare che se r_1 e r_2 sono uniformi in $[0,1]$ allora
 $x = \sqrt{-\log(r_1)} * \cos(2\pi r_2)$ è gaussiana (Metodo di Box-Muller)

Esercizio: campionare la distribuzione $P(x) = \exp\left(\frac{1}{1+x^2}\right) - 1$

Essendo $0 < t = \frac{1}{1+x^2} \leq 1$ si ha $\frac{e^t-1}{e-1} \leq t$ (Dimostratelo!) e quindi $P(x) \leq \frac{e-1}{1+x^2} = (e-1)g(x)$

Poiché $g(x)$ è normalizzabile e $P(x) \leq g(x)$ allora anche $P(x)$ è normalizzabile infatti

$$\int_{-\infty}^{+\infty} P(x) dx = k \approx 2.086...$$

k non può essere espresso in termini di funzioni standard ma può essere approssimato numericamente.

Quanto segue, però, non dipende da k e pertanto è inutile conoscerlo. Basta sincerarsi che esista finito!

Più utile invece è conoscere il massimo $P(0) = \max(P(x))$

Occorre normalizzare la distribuzione $g(x)$ e determinarne la cumulativa $G(x)$. Dall'integrale

$$\int_{-\infty}^x g(t) dt = \frac{\pi}{2} + \text{atan}(x)$$

si determina la cumulativa $G(x)$ correttamente normalizzata $\lim_{x \rightarrow \infty} G(x) = 1$

$$G(x) = \frac{1}{2} + \frac{1}{\pi} \text{atan}(x)$$

che è campionabile analiticamente. Se $r = \text{rand}()$

$$x = G^{-1}(r) = -\cot(\pi r)$$

Si accetta la x generando un secondo numero uniforme in $[0,1]$ e controllando se $r < \frac{P(x)}{P(0)}$

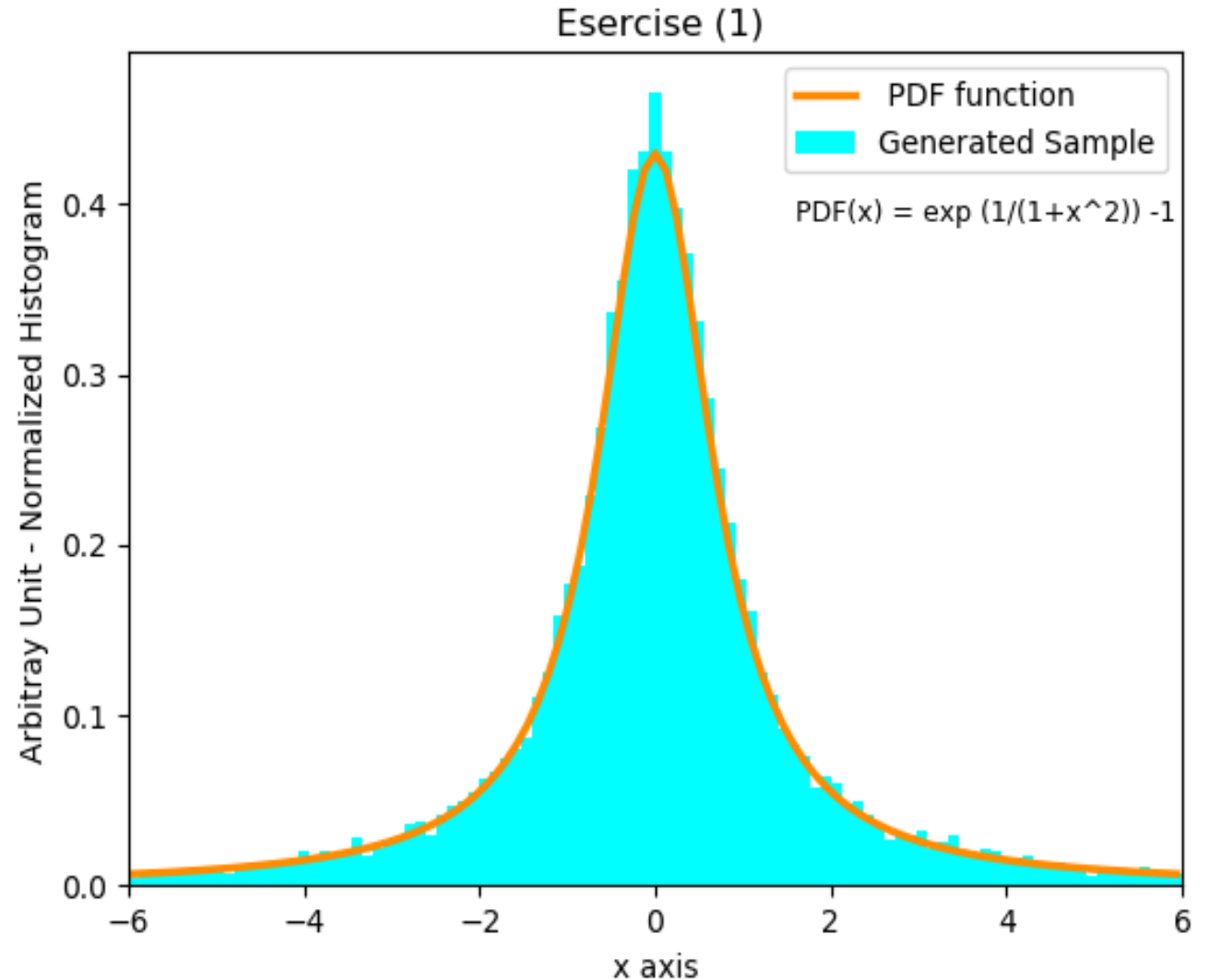
Esercizio 1

1) Define pdf (Probability Density Function)

```
def pdf(x):  
    y = np.exp(1/(1+x*x))-1  
    return y  
def g(x):  
    y = 1/(1+x*x)  
    y = y*2*pdf(0) # Maxima g(0) = pdf(0)  
    return y
```

2) Hit or Miss

```
while ipoint < npoint :  
    igen = -1 # igen = -1 rejection , igen=1 BINGO!  
    while igen < 0 :  
        r = np.random.uniform() #first random  
        t = np.tan(np.pi*(r - 0.5))  
        r = np.random.uniform() #second random  
        r=r*g(t)  
        if r < pdf(t) : # hit or miss  
            igen = 1  
    x[ipoint] = t  
    ipoint = ipoint+1
```



Esercizio 2

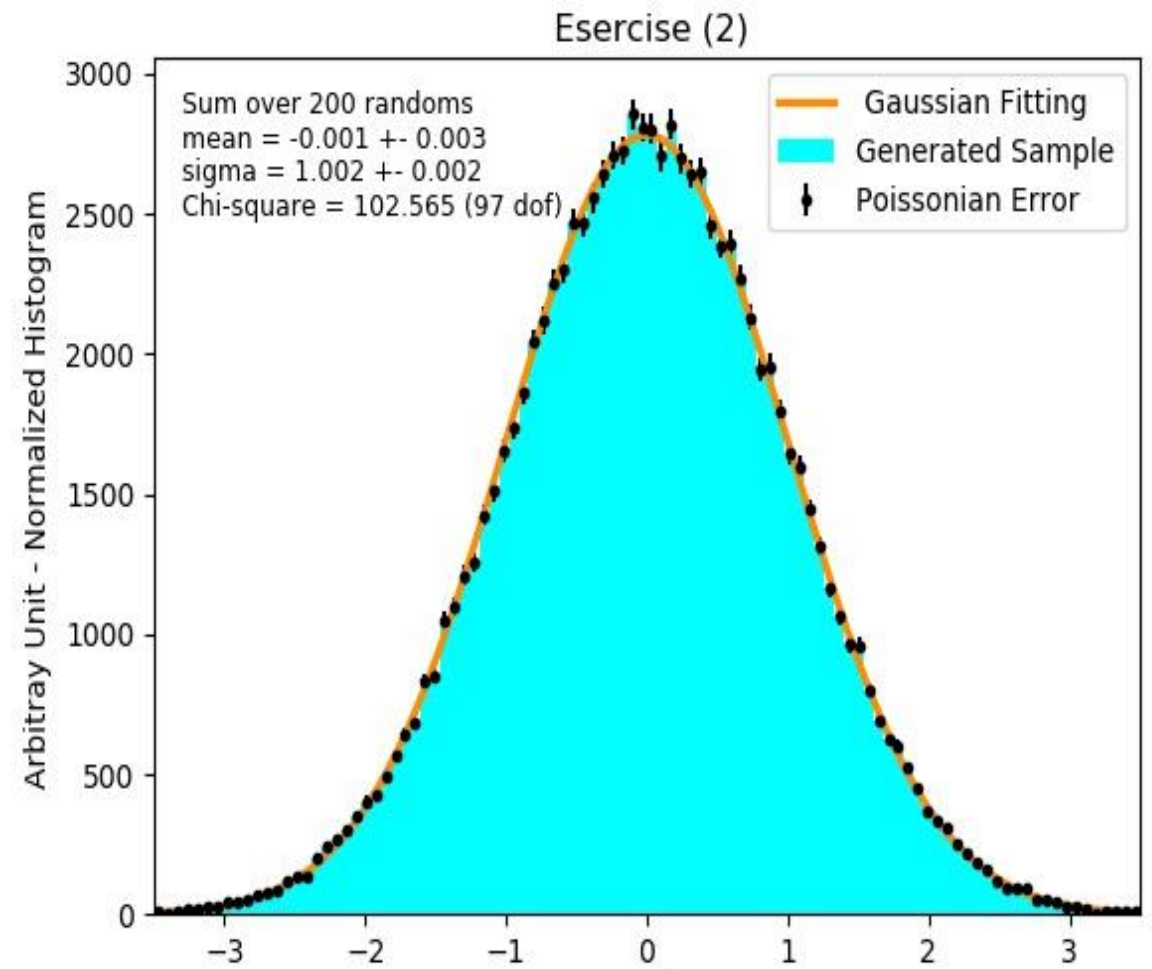
Se x_1, x_2, \dots, x_n sono variabili casuali (con distribuzione qualunque purché avente varianza definita) allora la loro media si distribuisce (asintoticamente per $n \rightarrow \infty$) come una gaussiana. (Teorema del limite centrale) Lo verifichiamo «computazionalmente» con x_1, x_2, \dots, x_n uniformi $[0,1]$

1) Initialization

```
nevent = 100000    #total number of trials
nd = 200           #number of randoms to be
                  #summed up in each trial
```

2) Running

```
ievent = 0
np.random.seed(314) # random seed
while ievent < nevent:
    array = np.random.random(nd)
    mu = np.mean(array) # mean
    std = np.std(array) # standard deviation
    std = std/np.sqrt(nd) # error on the mean
    z = (mu-1/2)/std # standardization
    x[ievent] = z
    ievent = ievent + 1
```

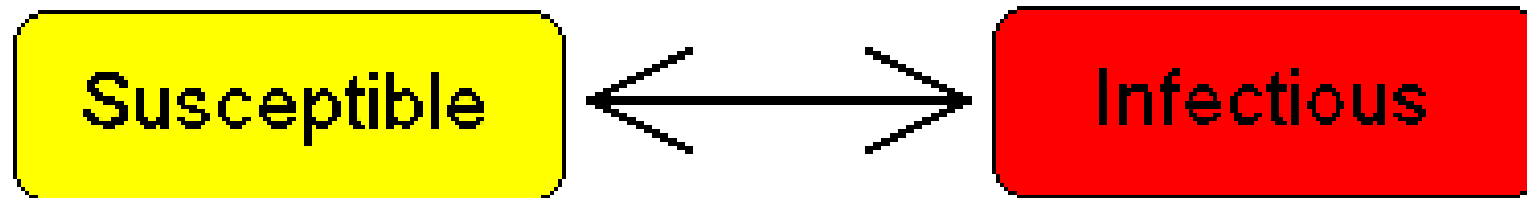


Alcune Applicazione dei Metodi Montecarlo

- Esempio: Modello Epidemiologico SI – studio degli effetti dovuti al tempo di latenza e di decorso della malattia
- Il «mistero» dei grandi chi-square :
- Valutazione dell'errore nel best fit di una circonferenza

Applicazione: Il modello epidemiologico SI

- Il più semplice modello che descrive l'evoluzione di una epidemia è il modello SI valido quando la malattia non dà immunizzazione. (ad esempio: il comune raffreddore o influenza!)
- La popolazione si divide in due categorie: «Suscettibili» (che possono infettarsi) e gli «Infetti». Se non c'è immunità, i guariti tornano nella categoria dei Suscettibili. In pratica si rimbalza da I a S



Equazioni del modello SI

$$\begin{cases} \frac{dS}{dt} = -\beta \frac{SI}{N} + \gamma I \\ \frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I \end{cases} \quad \longrightarrow \quad \begin{cases} \frac{dS}{dt} + \frac{dI}{dt} = 0 \\ \frac{dI}{dt} = aI(1 - bI) \end{cases}$$

- β e γ sono costanti dimensionate come l'inverso di un tempo
- β è il numero medio di contatti tra persone nell'unità di tempo moltiplicato per la probabilità che nel contatto avvenga il contagio di un suscettibile da parte di un infetto.
- Se il tempo necessario per guarire (cioè per tornare ad essere suscettibile dopo essere stato infettato) se dura D giorni allora $\gamma = 1/D$

Soluzione Logistica

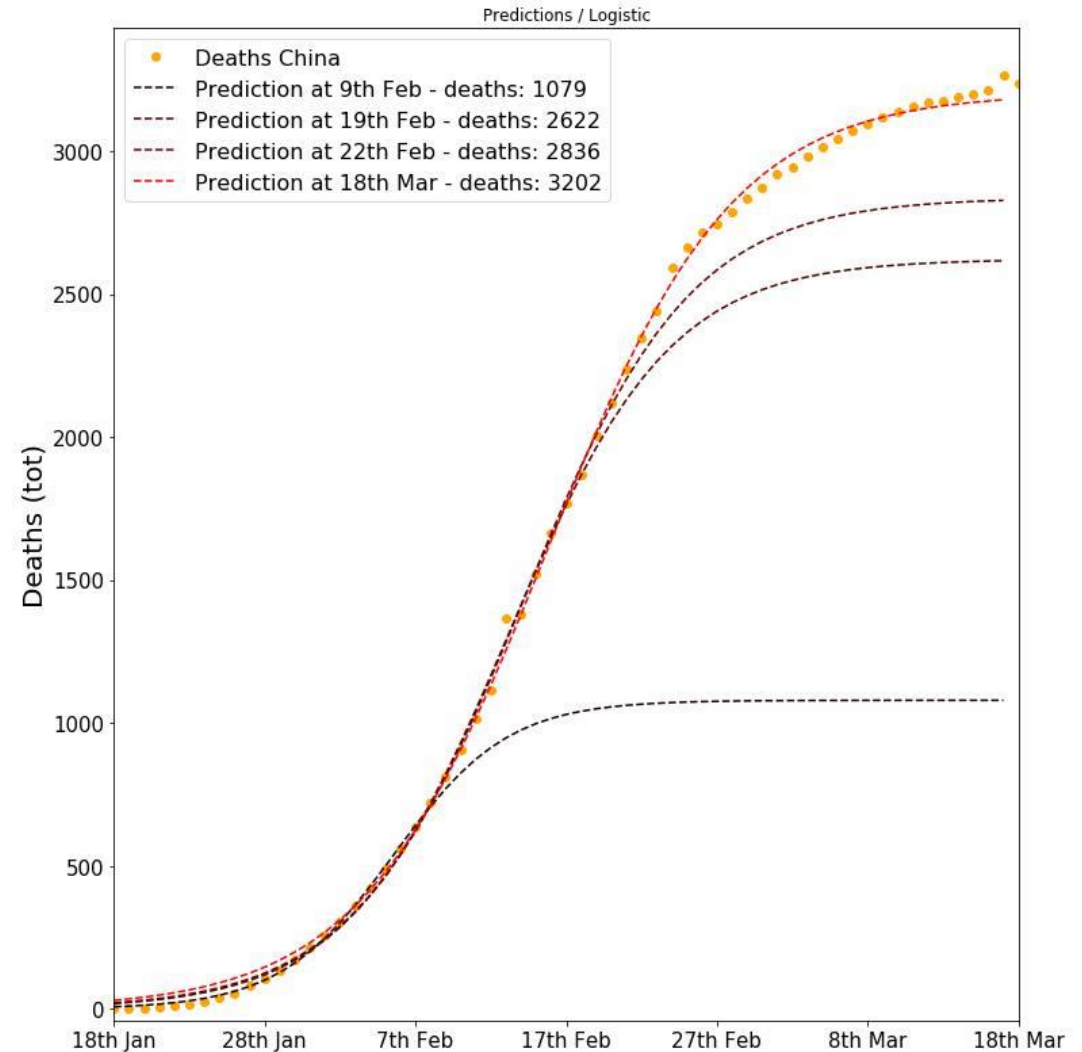
La soluzione per $I(t)$ è data dalla cumulativa della distribuzione «logistica»

$$L(t) = \frac{L}{2} (1 + \tanh(x)) \quad \text{con} \quad x = \frac{t - t_0}{\tau}$$

t_0 è il giorno di massimo contagio ovvero dove cade il massimo di dL/dt (punto di flesso per L)

τ è un tempo caratteristico del fenomeno.

I dati dalla Cina (dove l'epidemia è quasi conclusa) mostrano un **grossolano** accordo con il modello teorico. (Tutte le curve ad S si somigliano!!!!)



Campionare la distribuzione Logistica

CDF cumulativa $L(x) = \frac{1}{2} (1 + \tanh(x))$

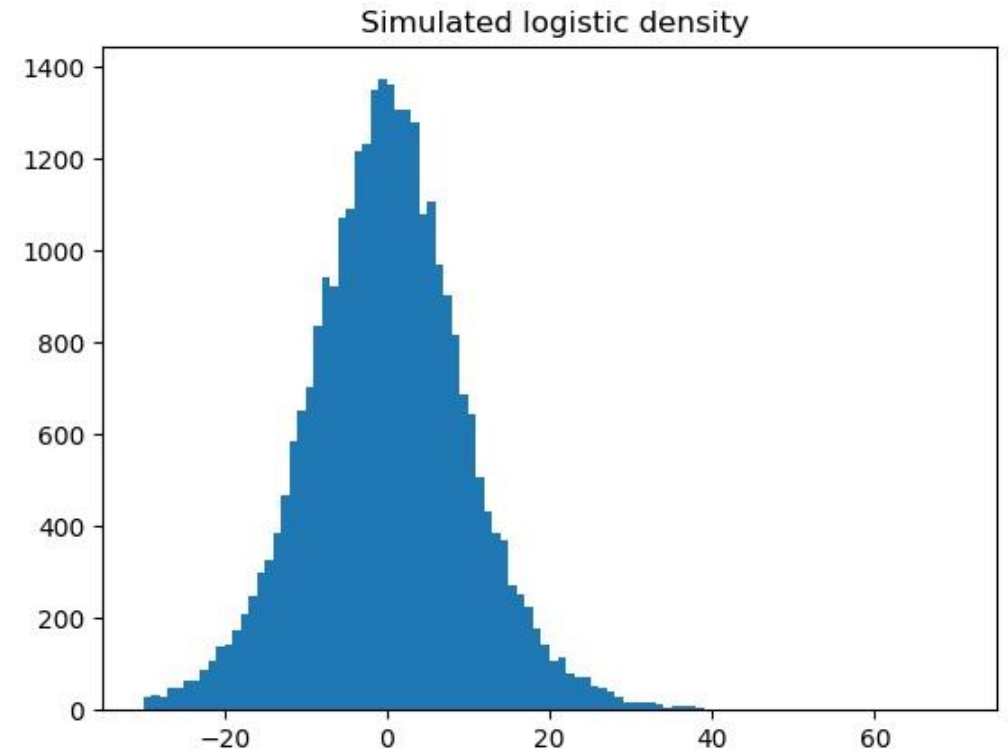
PDF densità $\frac{dL}{dx} = \frac{1}{2(\cosh(x))^2}$

$r = \text{random in } [0,1]$

$2r = 1 + \tanh(x)$

$x = \operatorname{atanh}(2r-1)$

```
while ipoint < npoint -1 :  
    r = np.random.uniform()  
    r = np.arctanh(2*r-1)  
    x[ipoint]= r*tau + t0
```



Effetto dovuto ai ritardi

- Mentre l'epidemia è in corso si vorrebbe «prevedere» quando e come si raggiungerà il picco per valutare se il sistema sanitario avrà posti a sufficienza per ospedalizzare chi ne ha bisogno
- $T(t)$ è la PDF del tempo che intercorre tra l'istante in cui si contrae l'infezione e la comparsa dei sintomi gravi che rendono necessario il ricovero
- La PDF dei malati gravi dP/dt (quanti ne arrivano tra t e $t+dt$) è una Logistica «ritardata» ottenuta convolvendo $dL(t)/dt$ con la distribuzione $T(t)$

$$dP(t)/dt = \int_0^\infty T(u) \frac{dL}{dx} (u - x) du \quad \text{dove} \quad x = \frac{t-t_0}{\tau}$$

Distribuzione dei Ritardi

- Una stima sui ritardi (tempi di incubazione, comparsa dei sintomi, etc...) nel caso del COVID19 è pubblicata in *Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections*, Journal of Clinical Medicine, 17/2/2020

Lognormal distribution	Incubation Period (days)	Onset to hospital admission (days)
Mean	5.0	10
Std	3	35

Inoltre il tempo di degenza e decorso della malattia valgono in media 13 +- 12 giorni nel caso di decesso e supera anche i 30 giorni nel caso di guarigione

Scelta del modello per i ritardo

Si assume un ritardo medio $\mu \approx 15$ giorni e deviazione standard $\sigma \approx 30$ giorni, distribuito secondo una Lognormale (suggerita dallo studio citato)

$$\Phi(x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{\log(x) - \mu}{\sigma\sqrt{2}} \right) \quad \text{CDF Cumulativa}$$

$$F(x) = \frac{d\Phi}{dx} = \frac{1}{x\sigma\sqrt{2\pi}} \exp \left[- \left(\frac{\log(x) - \mu}{\sigma\sqrt{2}} \right)^2 \right] \quad \text{PDF densità}$$

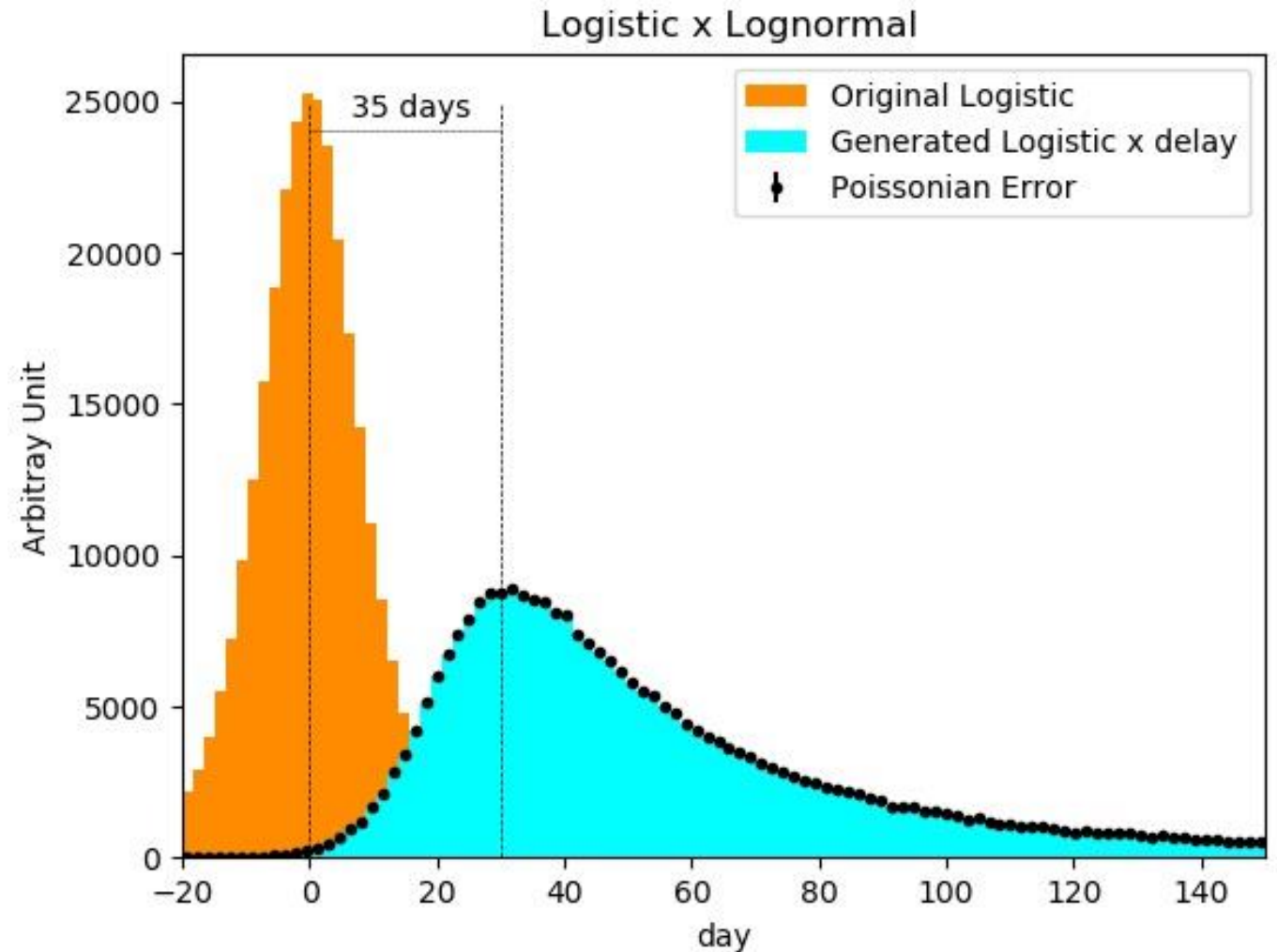
Simulazione dei ritardi

- L'integrale convolutorio Logistica x Lognormale non è analiticamente esprimibile con funzioni standard (Se volete divertirvi fatelo fare a Wolfram-Alfa!)
- In pratica si può fare con le tecniche di campionamento Montecarlo
 - 1) Si campiona la Logistica estraendo x
 - 2) Si campiona la Lognormale estraendo u
 - 3) Si «ritarda» $x = x + u$
 - 4) Si ripete molte volte per ottenere la distribuzione di x

Simulazione dei ritardi

```
# Generation parameters
tau = 10 #Logistic rising time
t0 = 0 # peak position
td = 15 # mean delay time
st = 30 # delay time standard deviation

while ipoint < npoint -1:
    r = np.random.uniform()
    r = np.arctanh(2*r-1)
    t = r*tau + t0
#Saving not yet delayed Logistic sampling
    xlog[ipoint] = t
# Lognormal delay inspired by
# https://www.mdpi.com/2077-0383/9/2/538
    r = np.random.lognormal(mean=0, sigma=1)
    delay = r*st+td
    x[ipoint] = t+delay
    ipoint = ipoint +1
```



Altri esempi di Integrali convolutori

L'esercizio basato sul modello SI serve per attirare l'attenzione sui meccanismi convolutori. Una applicazione meno drammatica può essere la risposta di un detector

Se un detector misura la variabile x con una risoluzione Gaussiana $G(x, \mu, \sigma)$ allora la distribuzione della x misurata differisce da quella delle x vere per effetto della convoluzione. (Trasformata di Weierstrass)

$$\left(\frac{dP}{dx}\right) \rightarrow \int_{-\infty}^{\infty} \left(\frac{dP}{dx}\right)_{x=u} G(u - x, \mu, \sigma) du$$

Il «mistero» dei grandi chi-square

Spesso il chi-square non torna con il valore atteso $\langle \chi^2 \rangle = \nu \pm \sqrt{2\nu}$

- Capita quando gli errori non sono gaussiani o lo sono ma sono stati stimati con poche misure (Il campione di misure raccolte è scarso e non ha ancora raggiunto la «centralizzazione» gaussiana degli errori nel senso del teorema del limite centrale)
- Capita quando si sottostimano o gli errori, se non si è tenuto conto di una sistematica o nel caso di sbagli grossolani e distrazioni varie
- Come esempio, simuliamo il caso della misura della accelerazione di gravità terrestre con un pendolo semplice. (tipico esercizio di LAB1)

Procedura per misurare g con un pendolo semplice

$$T = 2\pi \sqrt{\frac{L}{g}} \left(1 + \frac{\theta_0^2}{16} \right)$$

- Si scelgono nlen Lunghezze diverse
- Per ogni lunghezza si misura nmeas = 10 volte il periodo

Raccolta dati

- Per ogni lunghezza si ricava periodo medio ed errore
- Per ogni lunghezza si calcola il rapporto $\frac{T_i^2}{L_i} = s_i \pm e_i$ (i=1,...,nlen)
- Si determina la media \bar{s} (media pesata)

Algoritmo di Analisi

- Si determina $g = \frac{4\pi^2}{\bar{s}}$
- Si determina $\chi^2 = \sum_{i=1}^{nlen} \frac{(s_i - \bar{s})^2}{e_i^2}$

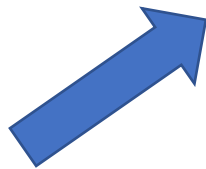
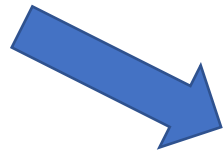
Output

Dati presi in
Laboratorio

Dati Simulati

Algoritmo di
Analisi

Output



Generazione dei dati simulati

La simulazione si effettua generando dati e «sporcandoli» (*to smear* in inglese!) con le distribuzioni degli errori che si suppongono note.

Esempio

- I periodi si «sporcano» per l'effetto delle fluttuazioni nel tempo di risposta umana (Dare lo start e stop al cronometro)

$T \rightarrow T + \Delta t * \text{random.normal}()$ Si assume smearing gaussiano

- Le lunghezze si «sporcano» con la risoluzione del righello

$L \rightarrow L + \Delta l * (2 * \text{random.uniform}() - 1)$ Si assume «smearing» uniforme

Si testa la bontà dell'algoritmo di analisi dandogli in ingresso i dati simulati e controllando lo scarto tra l'output e quello che ci aspetta teoricamente. Nel linguaggio della statistica: controllare la bontà di uno «stimatore».

IL metodo può inoltre suggerire la presenza di eventuali effetti sistematici e una cura per essi (Ad esempio: Occorre misurare l'ampiezza di oscillazione oppure basta accontentarsi che sia «genericamente» piccola ?)

Generazione dei dati simulati : ciclo sul numero di misure annidato nel ciclo sulle lunghezze. Ogni «smearing» corrisponde al campionamento di una distribuzione. L'angolo viene scelto a caso uniforme tra un minimo e un massimo. Lo sperimentatore non si preoccupa di misurarlo ! (Fa bene ?)

```
for ilen in range(len(lgen)):
    r = np.random.uniform()
    r = 2*r-1
    lrec[ilen] = lgen[ilen]+dl*r                                # smearing length
    for imeas in range(nmeas):
        a = 2*np.pi*np.sqrt(lgen[ilen]/g)
        r = np.random.uniform()
        theta = thetamin +(thetamax-thetamin)*r                #Initial angle smearing
        a = a*(1+theta**2/16)                                #Non perfect isochronism
        a = a+dt*np.random.normal()                            #Period smearing
        r = np.random.uniform()
        tgen[ilen][imeas] =a
```

Algoritmo di Analisi (o ricostruzione)

```
trec = np.mean(tgen, axis=1)
etrec = np.std(tgen,axis=1)/np.sqrt(nmeas) # error on mean
thetarec = (thetamax-thetamin)/2. #Correction for non isochronism
trec = trec/(1+thetarec*thetarec/16)

x = lrec

ex= np.ones(len(lgen), dtype=float)
ex = dl*ex/np.sqrt(12)
y,ey = trec**2, 2*etrec*trec

slope , eslope = y/x , slope*np.sqrt((ey/y)**2 + (ex/x)**2)
a ,b = sum(slope/(eslope**2)) , sum(1/(eslope**2))
slope_rec = a/b                #Best slope
eslope_rec = 1/np.sqrt(b)      # Error on best slope
a= (slope-slope_rec)/eslope    #Residuals

chisq , grec = sum(a*a) , 4*np.pi*np.pi/slope_rec
egrec = grec*eslope_rec/slope_rec    # Error on best g
```

Data reduction to mean periods and errors. Note the correction accounting for non-perfect isochronism, based on «average» oscillation angle!

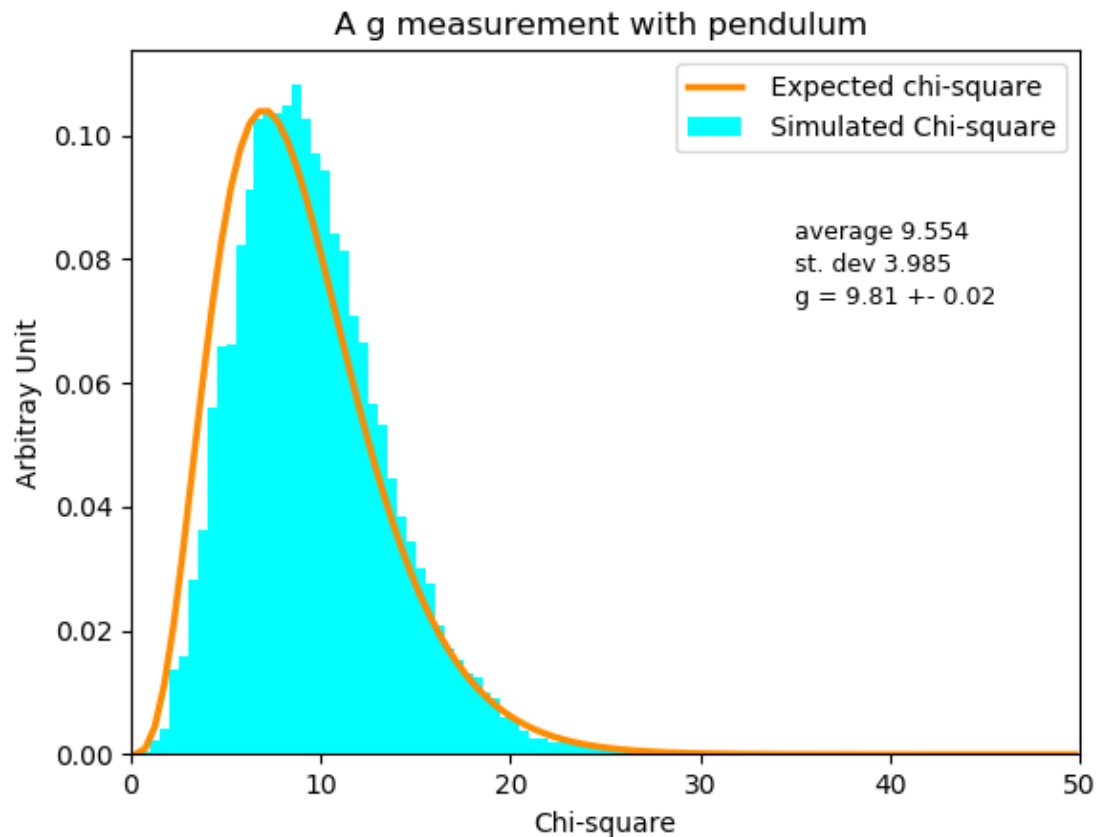
Error Propagation accounting for errors on y and x **as well** . Note the factor $\sqrt{12}$ Why is there?

Core of the analysis:

Equivalent to find the best fit to $y = m \cdot x$ where $m = \text{slope}$

Outputs : chi-square, measured g and error

Chi-square distribuzione nel caso «ideale»



- $N_{\text{meas}} = 10$
- Propagazione degli errori perfetta (su x e su y)
- Angoli piccoli $< 5^\circ$ ed effetto corretto

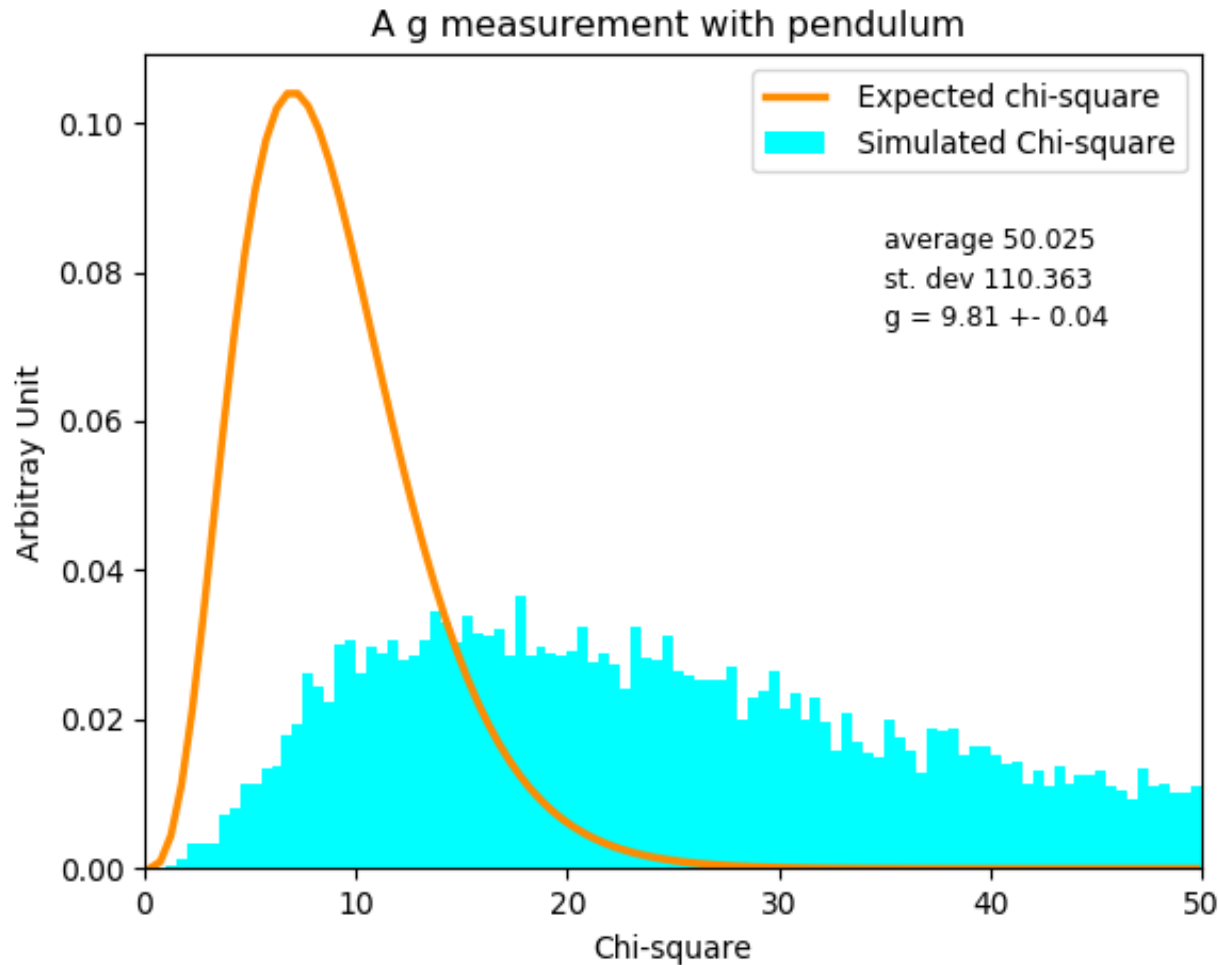
l'accelerazione ricostruita $g = 9.81 \pm 0.02$
torna entro gli errori con $g = 9.81$ usata in
generazione.

Il valor medio del chi-square 9.0 ± 4.2 è in accordo
con quella prevista teoricamente ($n.d.o.f = 9$)

Conclusione: L'algoritmo di Analisi sembra corretto

Nonostante tutto c'è una piccola discrepanza tra la
distribuzione teorica attesa e la distribuzione
ottenuta Perché ?

Chi-square distribuzione nel caso «lazy»



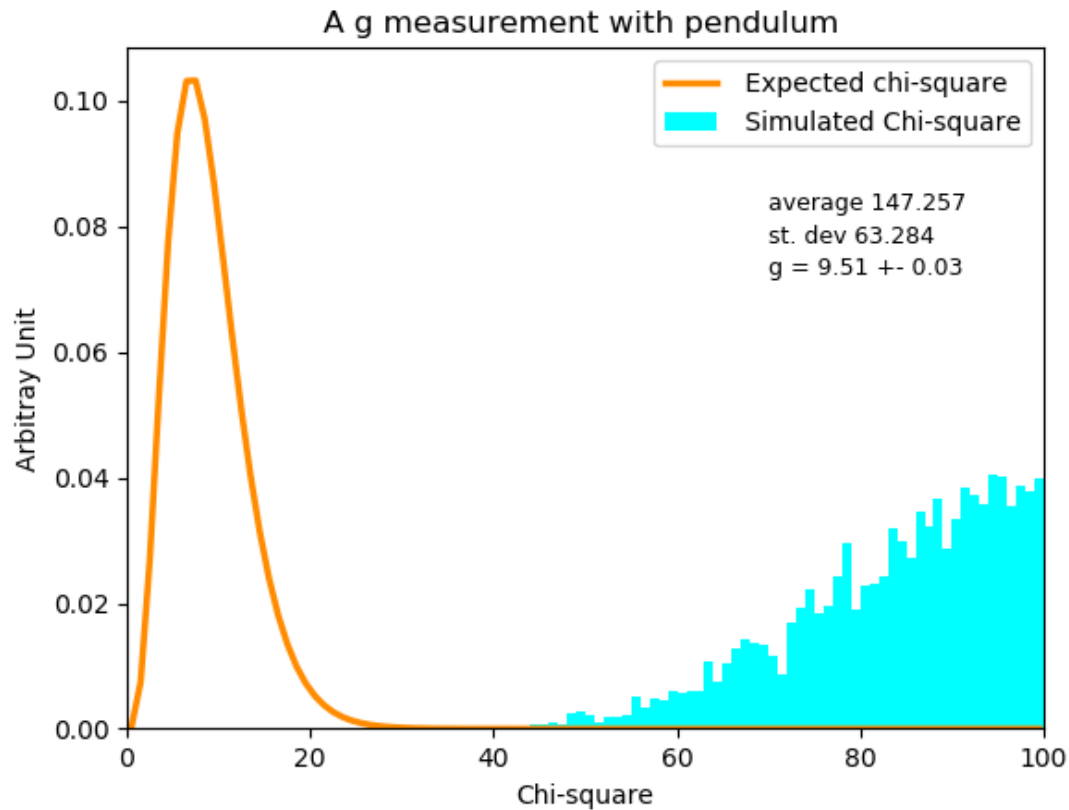
- $N_{\text{meas}} = 3$ **(Studente PIGRO che fa poche misure!!!)**
- Propagazione degli errori perfetta (su x e su y)
- Angoli piccoli $< 5^\circ$ ed effetto corretto

l'accelerazione ricostruita vale $g = 9.81 \pm 0.04$

La «pigrizia» ha fatto raddoppiare l'errore anche se non ha mosso il valore centrale

Il chi-square è totalmente fuori dalla previsione.
Perché?

Chi-square nel caso «inaccurate»

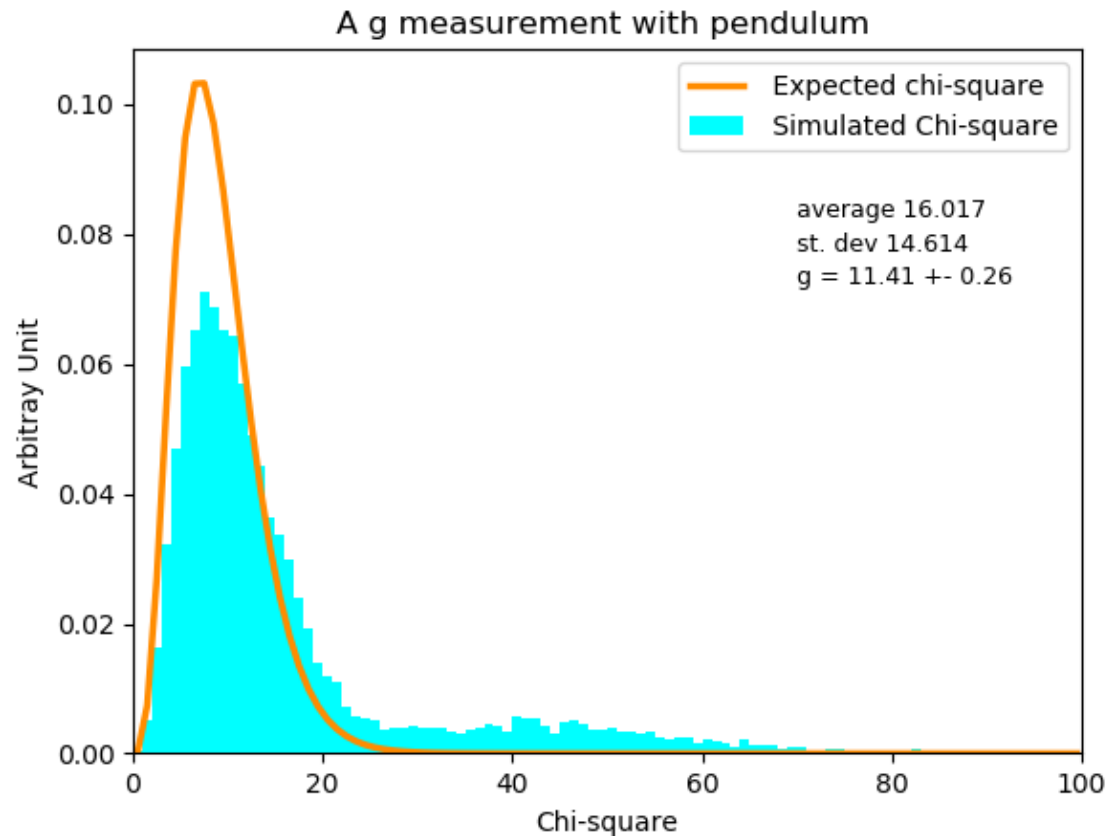


- $N_{\text{meas}} = 10$ Non è pigro ma non cura i dettagli
- Propagazione degli errori sbagliata (**non tiene conto dell'errore lungo x**)
- Angoli $< 30^\circ$ e non corregge per l'effetto
- Non include della lunghezza misurata il pesetto appeso

l'accelerazione ricostruita vale $g = 9.51 \pm 0.03$
L'inaccuratezza ha mosso il valore centrale di $10\text{-}\sigma$
rispetto al valore usato in generazione

Il chi-square è totalmente fuori scala !

Chi-square nel caso «trivial»



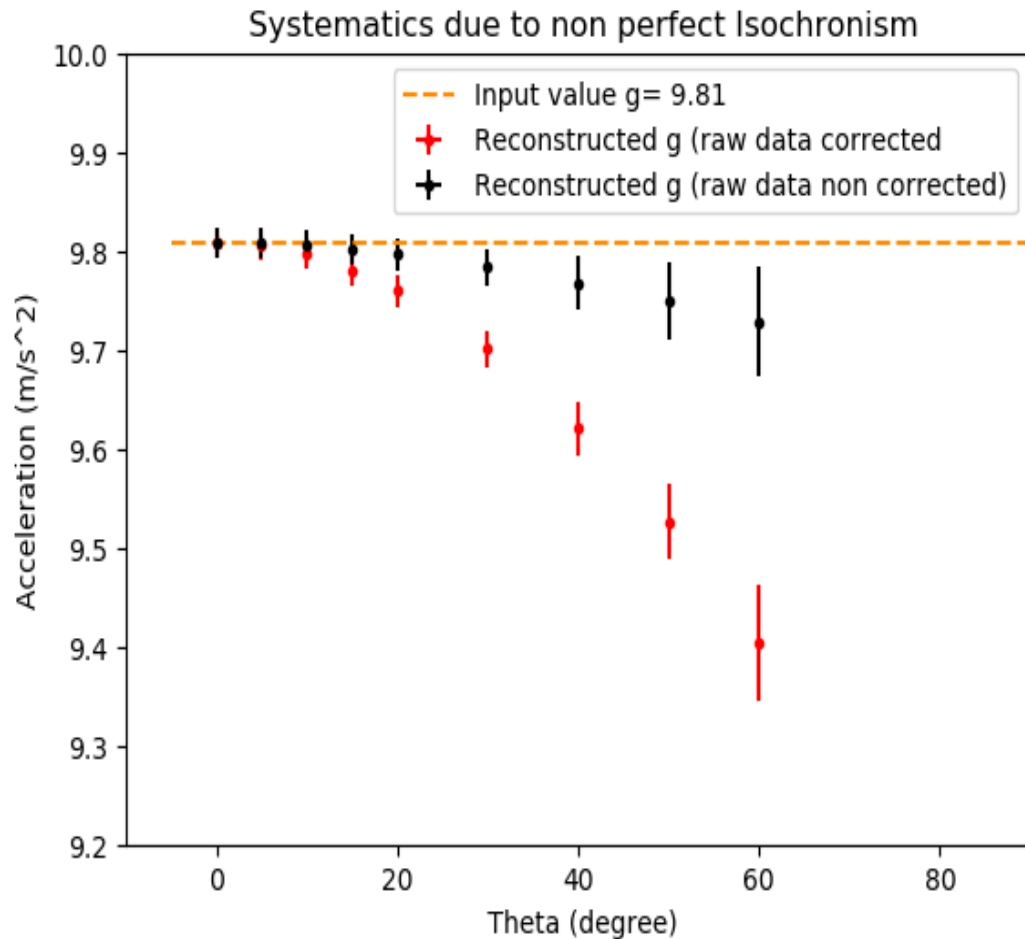
- $N_{\text{meas}} = 10$
- **Errore triviale.** In una certa frazione di casi crede di contare 10 oscillazioni ma in realtà sono 9 perché conta partendo da 1 e non da 0.

l'accelerazione ricostruita $g = 11.4 \pm 0.3$ è in buon accordo con quella sulla superficie di Nettuno!

Da notare l'insidioso chi-square che può anche dare valori ragionevoli.

Questo esempio dimostra che il chi-quadro come indicatore di qualità della misura qualche volta non funziona. La sua bontà non è condizione sufficiente a validare una misura!

Sistematica dovuta al non perfetto isocronismo



Possiamo decidere di non misurare l'ampiezza dell'oscillazione (risparmiamo sul goniometro) e sceglierla ogni volta a caso tra $[0, \theta]$ quindi applicare ai periodi ricostruiti una correzione media.

$$\theta_{\text{tarec}} = (\theta_{\text{amax}} - \theta_{\text{amin}}) / 2$$

$$t_{\text{rec}} = t_{\text{rec}} / (1 + \theta_{\text{tarec}} * \theta_{\text{tarec}} / 16)$$

Quale errore sistematico commettiamo se decidiamo di applicare la correzione media?

Errore sistematico e errore statistico sono paragonabili se $\theta < 15^\circ$

Esercizi

- Includere nella simulazione l'effetto dovuto alla non puntiformità

Suggerimento: Il momento di inerzia del peso non è in genere noto tuttavia è sicuramente della forma $I = k m L^2$ (dove L è la dimensione «tipica» del pesetto). Si trovi un upper limit per k . Si proponga una correzione ai raw data e si discuta l'effetto sistematico su g

- Includere nella simulazione l'effetto dovuto agli attriti

Suggerimento: Informarsi su come è definito il «fattore di qualità» di un pendolo. Includerlo nella simulazione. Dire come e con quale precisione si può misurare. Dire quali sono i suoi effetti sistematici sulla misura di g .

Best Fit a una circonferenza



Il rapporto tra raggio esterno e raggio interno di un doppio arcobaleno è legato all'indice di rifrazione dell'acqua

$$R1/R2 = f(n)$$

Idea sperimentale: si misurano raggi e si determina n invertendo $f(n)$

Come si determina il best fit alla circonferenza?

Quale algoritmo di ricostruzione va usato ?

Algoritmo ingenuo !

- minimizzare

$$\chi^2 = \sum \frac{(y_i - f(x_i))^2}{\sigma^2 + (f'(x_i)\sigma)^2}$$

$$f(x) = y_c \pm \sqrt{R^2 - (x - x_c)^2}$$

Ci sono due problemi

- Il doppio segno (facile da risolvere!)
- Se $x \approx x_c$ la derivata diverge!

Algoritmo meno ingenuo

- Per trovare la migliore circonferenza compatibile con i punti (x_i, y_i) $i=1, \dots, n$ ($n>3$) conviene
- effettuare la trasformazione che porta il baricentro nell'origine
 $u=x-x_b$; $v=y-y_b$
- Quindi minimizzare la quantità «chi-square-like»

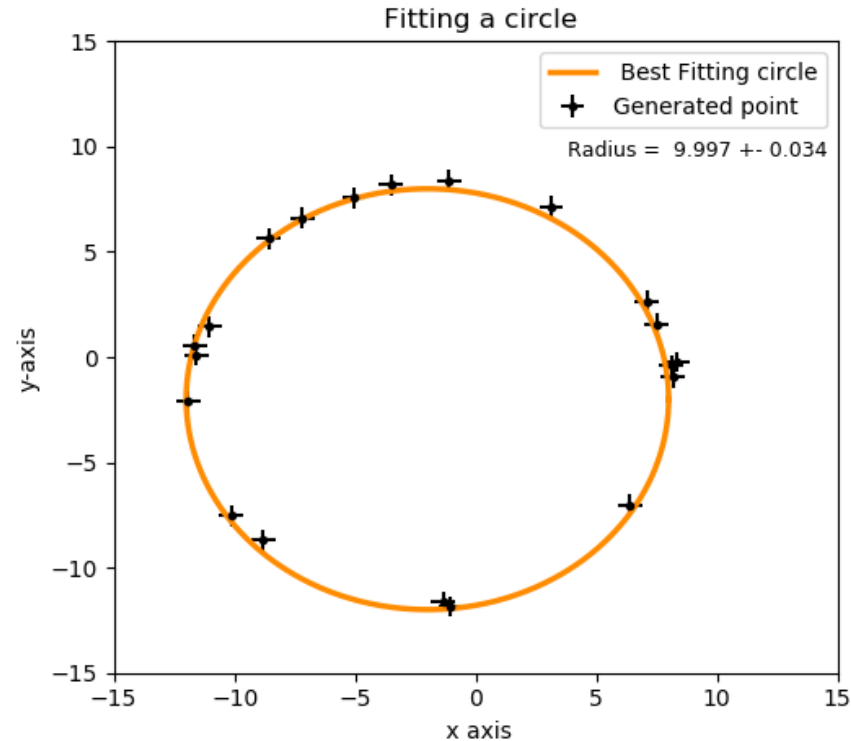
$$\chi^2 = \sum (g(u_i, v_i) - 0)^2$$

$$g(u, v) = R^2 - (u - u_c)^2 - (v - v_c)^2$$

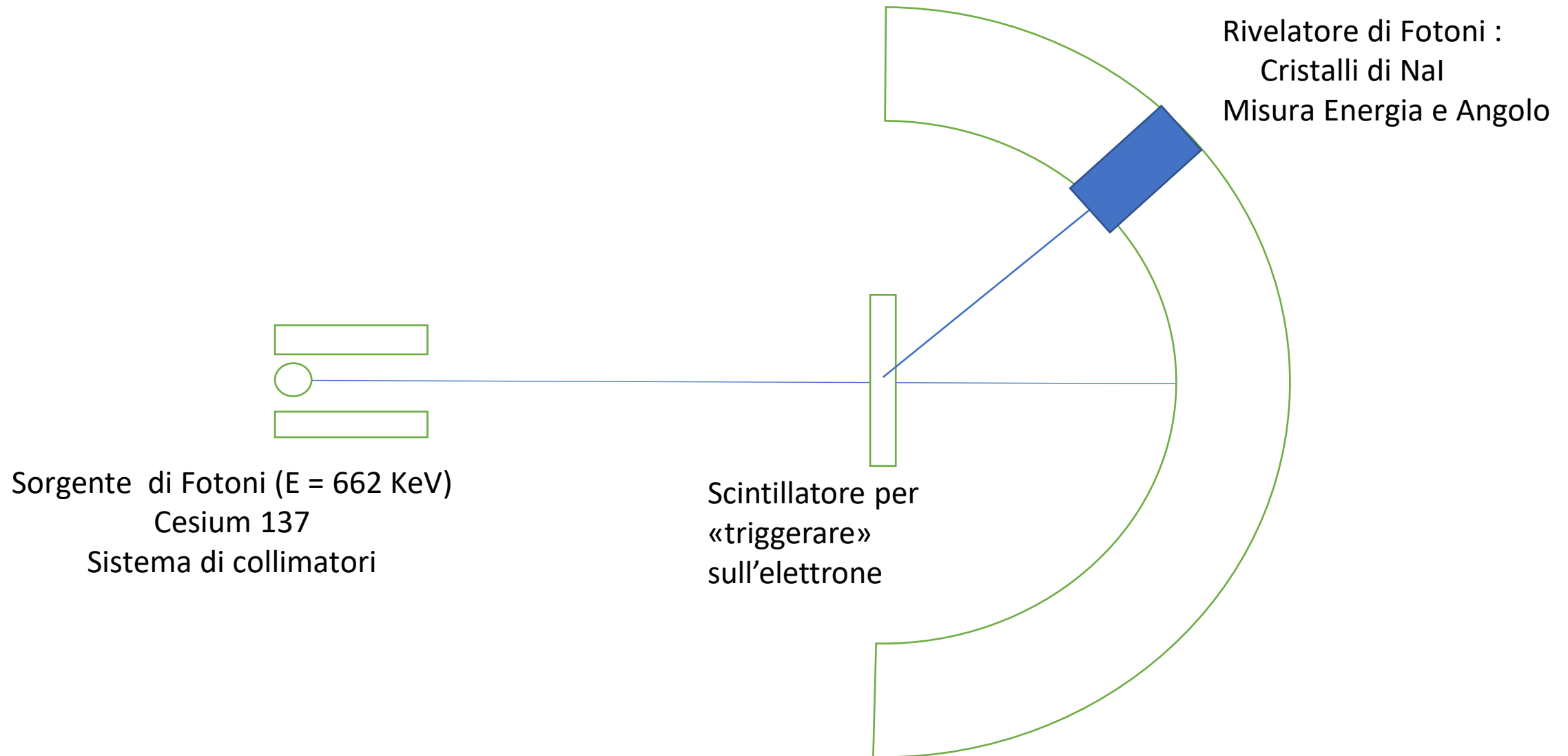
Esercizio Risolvere analiticamente la minimizzazione e controllare con quella implementata nel codice circle.py

La simulazione mette in evidenza i dettagli

- Algoritmo di fitting ricostruisce bene il raggio ? Con quale errore ?
- Quale è la correlazione tra gli errori sul Raggio e quelli sulle Coordinate del centro ?
- Come scala l'errore con il numero di punti ?
- Quale è l'effetto di poter usare solo «mezza» circonferenza ?
- Se vogliamo misurare l'indice di rifrazione con un errore del 0.1% di quanti punti abbiamo bisogno?
- **N.B. Il metodo è sensibile all'indice di rifrazione n oppure a $(n-1)$?**



Effetto Compton e Misura della massa dell'elettrone

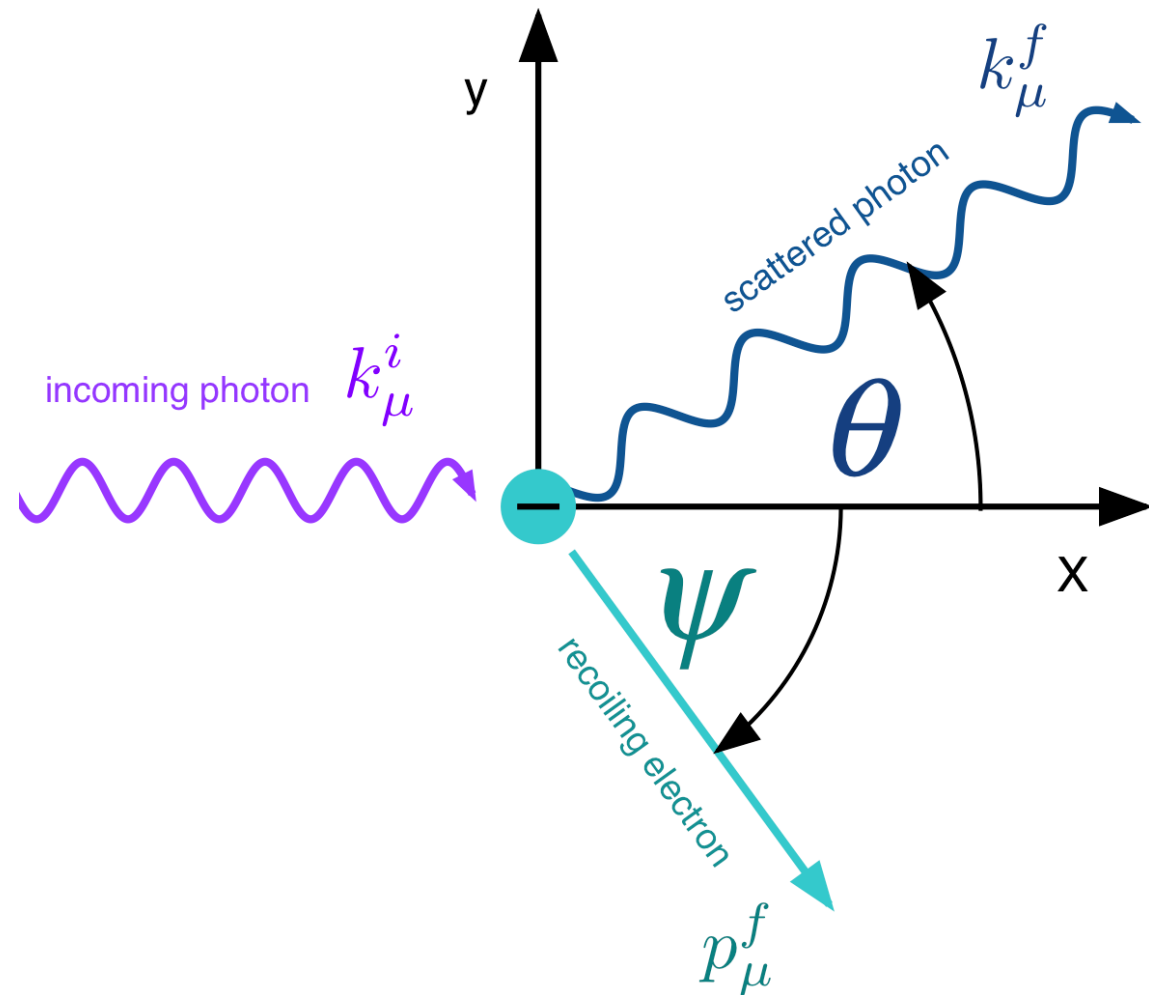


Cinematica $\gamma + e^- \rightarrow \gamma + e^-$

$$x = 1 + \varepsilon \sin^2\left(\frac{\theta}{2}\right)$$

$$1 < x = \frac{k}{k_f} < 1 + \varepsilon$$

$$\varepsilon = \frac{2k}{m_e}$$



Dinamica $\gamma + e^- \rightarrow \gamma + e^-$

- **basse energie ($k \ll m_e$)** il fotone «rimbalza» coerentemente sull'elettrone (Lo vede come se fosse un centro di scattering rigido di massa infinita...) e mantiene la propria energia $k \approx k'$ Il fotone emerge con distribuzione angolare

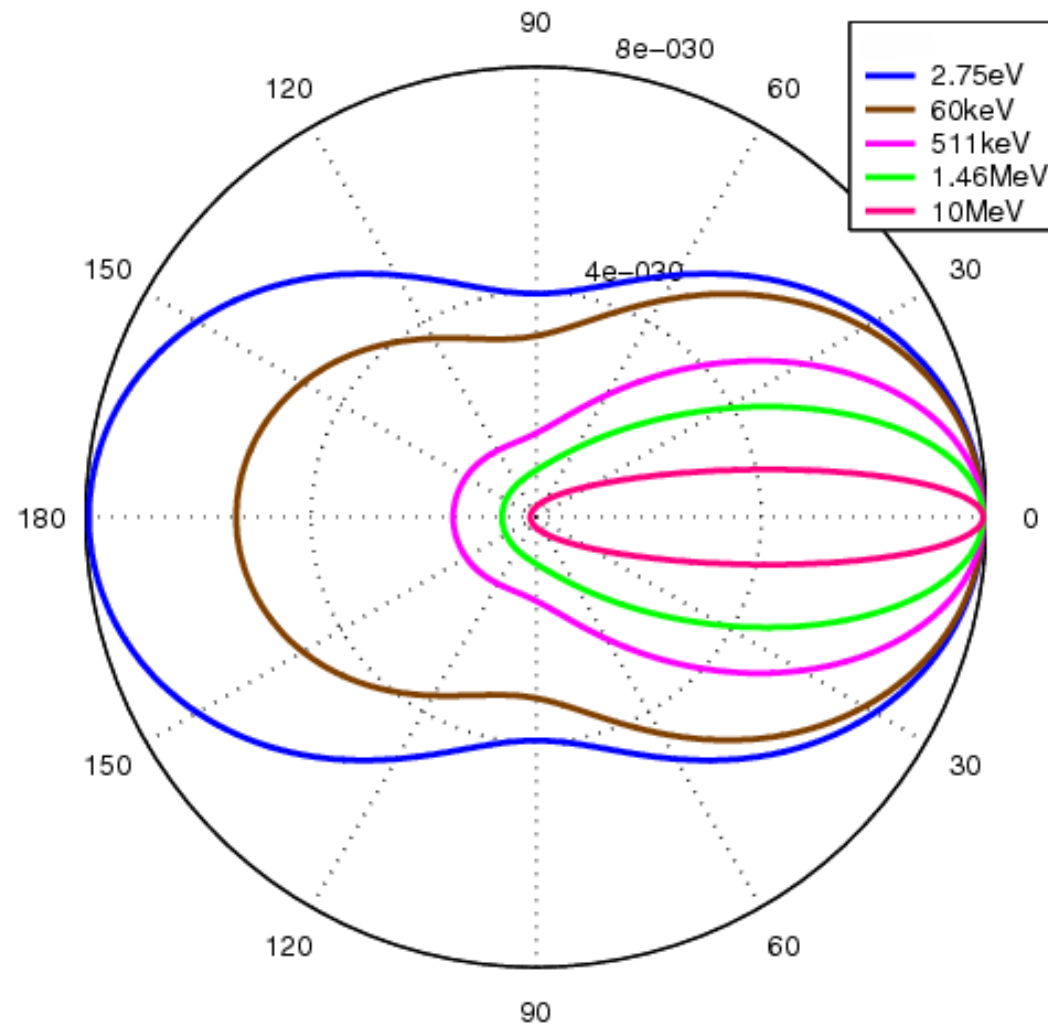
$$\frac{d\sigma}{d\cos\theta} = \frac{1}{2} r_e^2 (1 + \cos^2\theta) \quad (\text{Thomson})$$

- **Alte Energie** la sezione d'urto diventa meno efficiente

$$\frac{d\sigma}{d\cos\theta} = \frac{1}{2} r_e^2 \left(\frac{k_f}{k} \right)^2 \left(\frac{k_f}{k} + \frac{k}{k_f} - \sin^2\theta \right) \quad (\text{Klein-Nishina})$$

Scattering Compton

Diagramma polare
Perdita di simmetria
alle alte energia



Esercizi

- Dimostrare che a basse energie Klein-Nishina si riduce a Thomson
- Trovare l'energia massima trasferita all'elettrone
- Studiare la asimmetria tra scattering in avanti e scattering indietro
- Dimostrare che sul piano (x,y) lo Scattering equivale alla matrice

$$S = \frac{k_f}{k} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \quad \text{quindi} \quad \begin{pmatrix} E' \\ p'_x \\ p'_y \end{pmatrix} = S \begin{pmatrix} E \\ p_x \\ p_y \end{pmatrix}$$

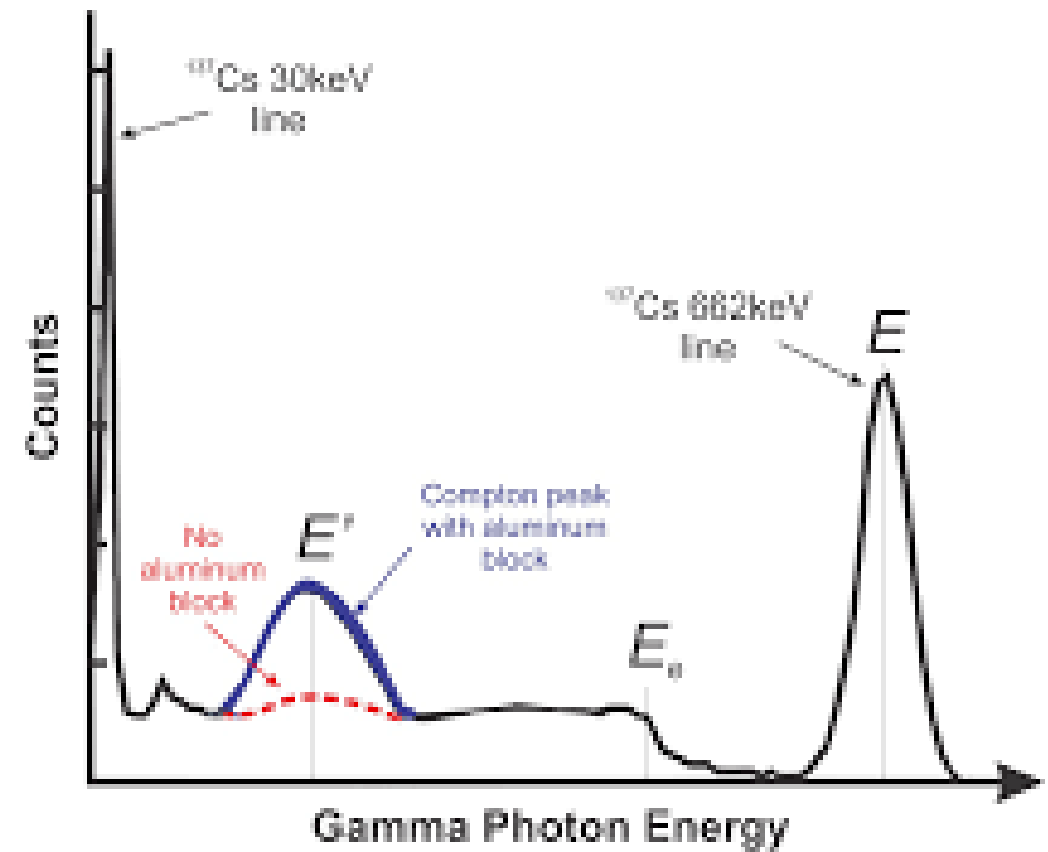
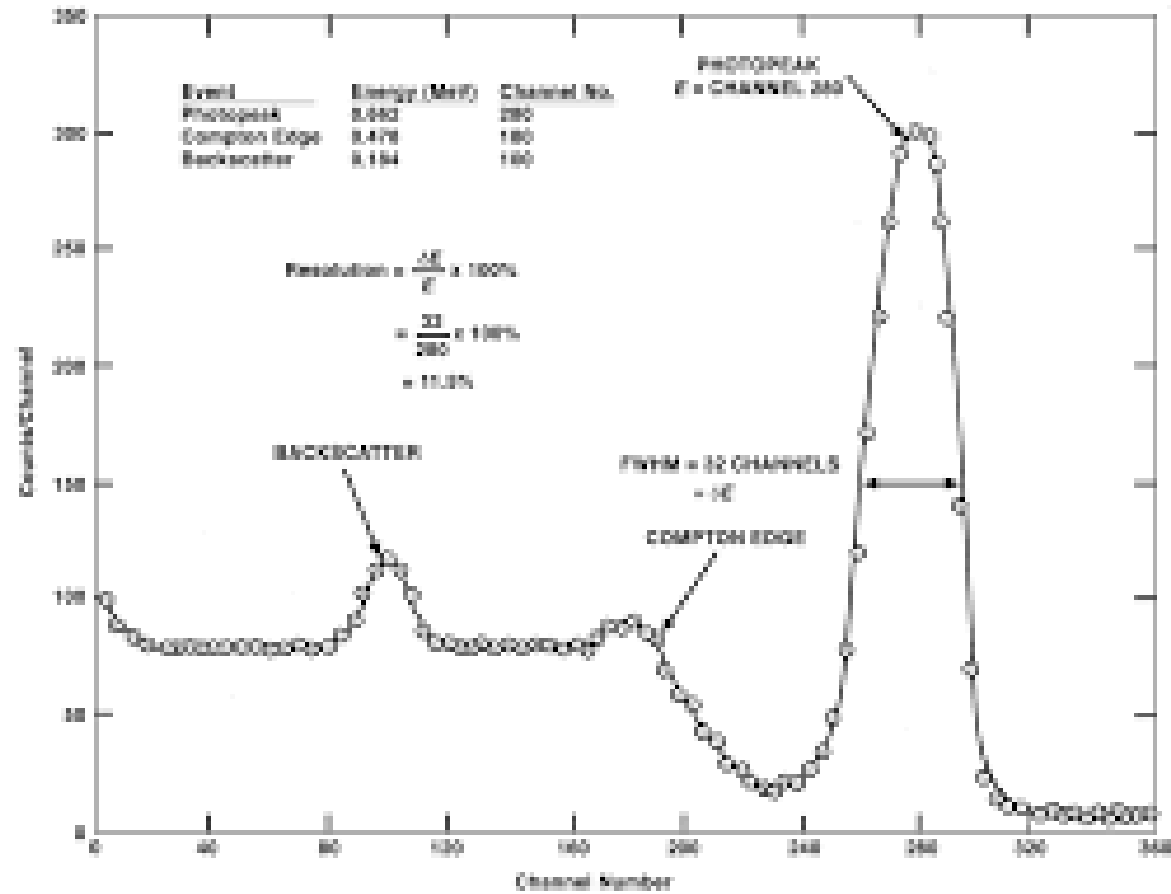
Cosa può dire la simulazione

- Usando fotoni da 662 KeV (cesium 137) si può misurare la massa dell'elettrone? (Come ? Con quale accuratezza e precisione?).
- Si «vedono» gli effetti relativistici sulla sezione d'urto differenziale ?
- Ci sono effetti dovuti al bordo del collimatore ?
- «Triggerare» sul segnale dell'elettrone introduce distorsioni? (Quali ?)
- Quando il fotone incide sul blocco di NaI può fare back-scattering. Cosa succede alla energia misurata ? Cercare informazioni sulla «spalla Compton» (Compton Edge!) e se qualcuno lo ha già simulato nel caso «rubare» la simulazione... (**Montecarlo Condensato**)

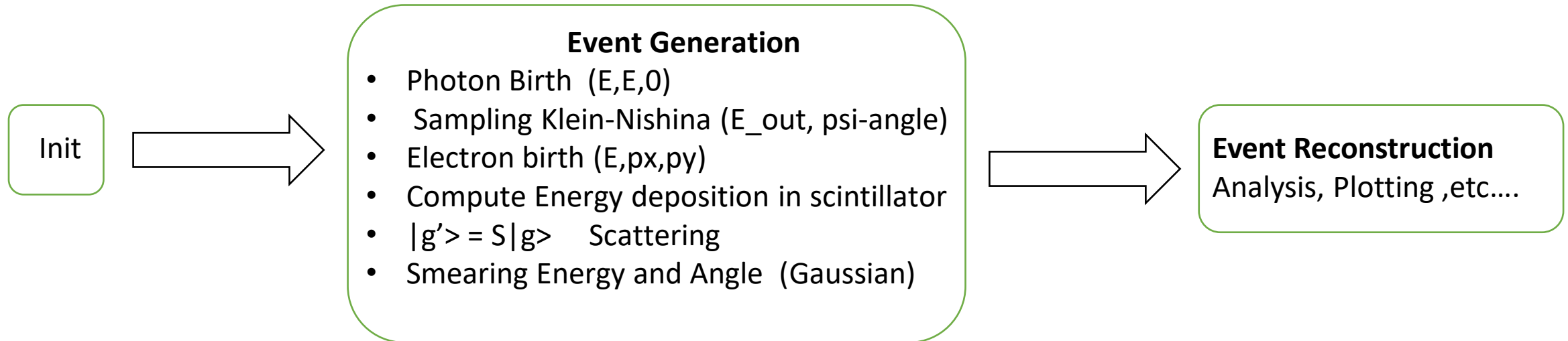
Problemi nella Simulazione

- Trovare un modo efficiente per campionare la Klein-Nishina (proporre metodi) Alcuni usano la variabile $y = k/k'$
- Risoluzione dei rivelatore $\frac{dE}{E} = \frac{a}{\sqrt{E}}$ (sensata ?)
- Come si simula il segnale dell'elettrone nello scintillatore ? (Dove si mette la soglia)

Compton Edge



Codice compton.py



Particelle cariche nella Materia

Particelle cariche (Elettroni, protoni, alfa) sono soggette alle interazioni con gli atomi del materiale attraversato. I fenomeni tipici sono:

- Ionizzazione (Produzione di raggi delta)
- Eccitazione degli atomi
- Radiazione di frenamento
- Effetto Cherenkov
- Radiazione di transizione (si verifica alle giunzioni)
- Interazioni nucleari

La particella perde energia lungo il cammino, fa scattering Multiplo (randomizzazione della direzione del momento) e infine si ferma ed è «assorbita» nel materiale.

Particelle cariche Pesanti

- Se $M_X \gg m_e$ i processi dominanti sono
- Collisioni elastiche nel campo Coulombiano dei nuclei
 $X + N \rightarrow X + N$ (Scattering, randomizzazione della direzione)
- Collisioni Inelastiche Coulombiane sugli atomi
- $X + A \rightarrow X + A^* \rightarrow X + A + \gamma$ (eccitazione EM dell'atomo)
- $X + A \rightarrow X + A^+ + e^-$ (Ionizzazione, produzione di raggi delta)
- Trascuriamo le interazioni Nucleari Ovvero quelle che «spezzano» il nucleo atomico

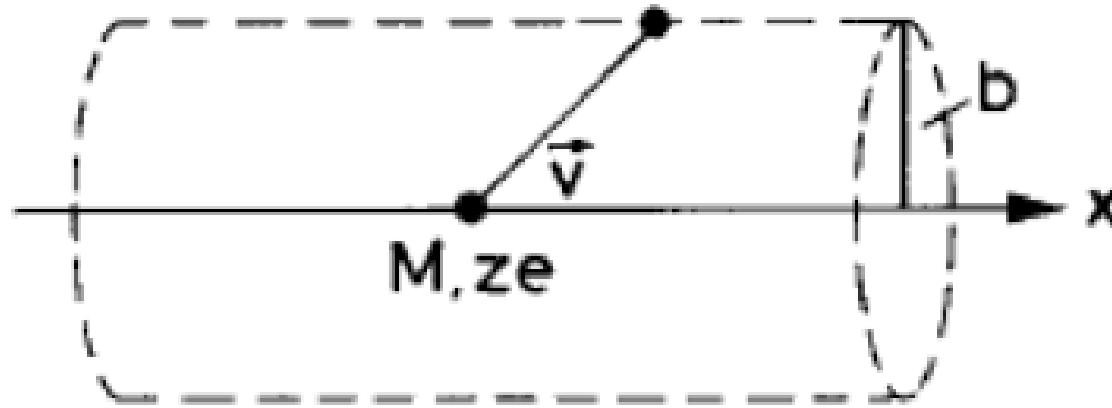
Perdita di energia per ionizzazione

- Se $M_x \gg m_e$ e $p \approx 1 \text{ GeV}$ allora domina l'interazione elettromagnetica e l'energia persa per ionizzazione.

Ricordiamo la trattazione «classica» data da Bohr (1913) [Che aiuta a stabilire gli ordini di grandezza, costanti coinvolte, etc..]

La trattazione QM si deve a Bethe-Block (1930) e citeremo solo il risultato.

Energia persa per ionizzazione nell'unità di cammino dE/dx (Bohr 1913)



Idea del calcolo : Considerare una particella carica che si muove lungo l'asse x (circondata dal materiale) e calcolare il «momento trasferito» a tutti gli elettroni che stanno nel guscio cilindrico $2\pi b db dx$ quindi integrare su ragionevoli estremi $b_{\min} < b < b_{\max}$ Il tratto dx è scelto abbastanza piccolo in modo che la velocità in esso si possa considerare costante!!!

Assomiglia tanto ad uno «splash»

- Il calcolo si basa sul Teorema dell'Impulso $F dt = dp$ (Secondo principio della dinamica «classica»)
- Potete immaginarlo come un esercizio di Fisica 1 in cui si chiede: Una palla da Bowling attraversa uno strato dx uniformemente riempito con palline da ping pong. Calcolare l'energia persa dalla palla da bowling nell'attraversare lo strato. Dire se e dove eventualmente si ferma se lo strato è abbastanza spesso.
- Oppure: un tuffatore si lancia da un trampolino di altezza h.. Dire a quale profondità dal pelo dell'acqua si ferma (e Archimede lo riporta su!!!)
- Oppure un meteorite si schianta sulla Luna.. Stimare la profondità del cratere
- Immaginare altri casi di «Splash» 😊

Consideriamo $X = \text{protone}$ ($z=1$ e $q=e$)

- Il momento trasferito all'elettrone che si trova in ($r=b$ e $x=0$) è ortogonale all'asse x (Momento trasverso solo!) (Dire perché quello longitudinale è nullo)

$$\Delta p = \int_{-\infty}^{+\infty} F dt = \int_{-\infty}^{+\infty} F \frac{dx}{v} = \frac{q^2}{v} \int_{-\infty}^{+\infty} \frac{dx}{(x^2 + b^2)^{3/2}} = \frac{2q^2}{bv}$$

L'energia persa sul singolo elettrone vale

$$\Delta E = -\frac{\Delta^2 p}{2m_e} = -\frac{4q^4}{b^2 m_e v^2}$$

va «integrata» su tutti gli elettroni presenti nel guscio cilindrico. Quanti ce ne sono nel tratto dx ?

$$d^2 N = 2\pi \rho b db dx$$

(ρ = densità di elettroni, quanti ce ne sono per m^3)

E ora.... Integrare su tutto

$$\frac{dE}{dx} = -\frac{4\pi\rho q^4}{m_e v^2} \int_{b_{min}}^{b_{max}} \frac{db}{b} = \frac{4\pi\rho q^4}{m_e v^2} \log\left(\frac{b_{max}}{b_{min}}\right)$$

Come scala lo spazio di frenata con l'energia iniziale?

Confrontare la profondità (range) raggiunta da un protone da 1 MeV e uno da 100 MeV (Energia cinetica) in acqua

Domanda: Perché abbiamo trascurato la perdita di energia dovuta alle collisioni Coulombiane con i nuclei atomici ? Quanto è più piccola rispetto a quella dovuta agli elettroni? Quale effetto provoca sul Range?

Domanda: Cosa diavolo sono b_{min} e b_{max} ?

Scelta di b_{min} e b_{max} ... («Black Magic» !!!)

- Parametro di impatto minimo perché non ha senso scendere al di sotto della lunghezza d'onda di De Broglie dell'elettrone (visto dalla particella carica incidente)
- Parametro di impatto massimo: Il tempo di interazione (come si valuta?) deve essere più breve del «periodo» dell'elettrone per garantire un «efficace trasferimento di energia»

$$b_{min} \approx \frac{h}{p} = \frac{h}{\gamma m_e v} \qquad b_{max} \approx \frac{\gamma v}{\langle v_e \rangle}$$

La «frequenza media» si stima attraverso il potenziale di ionizzazione $I = h \nu$

In conclusione

- $\frac{dE}{dx} = - \frac{4\pi\rho q^4}{m_e v^2} \log \left(\frac{\gamma^2 \beta^2 m_e c^2}{I} \right)$ (Bohr 1913)
- $\frac{dE}{dx} = - \frac{4\pi\rho q^4}{m_e v^2} \left[\frac{1}{2} \log \left(\frac{\gamma^2 \beta^2 m_e c^2 T_{max}}{I^2} \right) - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right]$ (B.B. 1930)
- $T_{max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma \frac{m_e}{M} + \frac{m_e^2}{M^2}}$
- Energia Massima trasferibile al singolo elettrone (Raggio delta)

Esercizio: Progettare un esperimento per decidere se tutto quanto detto ha riscontro in natura